Gene cluster involved in nogalamycin biosynthesis, and its use in production of hybrid antibiotics

Field of the invention

This invention relates to the gene cluster for nogalamycin biosynthesis derived from *Streptomyces nogalater*, and the use of the genes therein to obtain novel hybrid antibiotics for drug screening.

Background of the invention

Anthracyclines are antitumor antibiotics, mainly produced by *Streptomyces* sp. Daunomycin family of anthracyclines is commercially most important, since almost all of the around ten anthracyclines currently in clinical use, or in late clinical trials for cytotoxic drugs, belong to this family. Despite the long history of anthracyclines, three decades or so, the studies on their biosynthesis are still going on, and there is further interest to obtain novel molecules for the development of cancer chemotherapeutics. A method currently used for finding novel molecules for drug screening is genetic engineering. Cloning the genes for anthracycline biosynthesis facilitates the production of hybrid anthracyclines, as well as their use in combinatorial biosynthesis to generate novel molecules.

Nogalamycin, which was first described by Bhuyan and Dietz in 1965, is an anthracycline antibiotic produced by *Streptomyces nogalater*. It is highly active against tumor cells, whereas toxic properties of this compound have prevented its progress to clinical trials (Bhuyan and Smith, 1975). However, menogaril (7–O–methylnogarol) is a semisynthetic derivative of nogalamycin, and its value in the treatment of cancer has been studied (e.g. Yoshida *et al.*, 1996), the interest being now mainly in Japan. Structurally nogalamycin (Fig. 1) differs from most other anthracyclines, as e.g. from the daunomycin family, in two noteworthy features: (i) The stereochemistry at position nine is opposite, and (ii) it has a sugar moiety, in which nogalamine is attached at position 1 by a typical glycosidic bond, but it is also attached to carbon 2 by an

extraordinary C–C bond. Structural elucidation of nogalamycin was reported by Wiley *et al.* (1977). Furthermore, biosynthetic studies of nogalamycin have been published by Wiley *et al.* in 1978 giving information of the building blocks: The aglycone moiety is built from ten acetates; the neutral sugar, nogalose, is derived from glucose; and methyl groups of both of the sugars, nogalamine and nogalose, are transferred from methionine. The origin of nogalamine was not clearly solved by Wiley, but most probably nogalamine is also derived from glucose.

Molecular cloning of biosynthesis genes for anthracyclines has facilitated the studies on molecular genetics, providing tools for rational modifications of the structures, while also for surprising combinations with other antibiotics. Most of the interest has focused on daunomycin biosynthesis genes, as reported in several publications (Lomovskaya *et al.*, 1998; Rajgarhia and Strohl, 1997 and references therein). Some genes for aclacinomycin biosynthesis from *S. galilaeus* (Fujii and Ebizuka, 1997) and for rhodomycin biosynthesis from *S. purpurascens* (Niemi *et al.*, 1994) have been cloned as well. We have cloned the biosynthesis genes for nogalamycin, and successfully used the genes for producing hybrid anthracyclines. Most of the genes are involved in polyketide pathway, being responsible for the formation of a tricyclic intermediate, and they are reported in Ylihonko *et al.*, 1996a and b, and by Torkkell *et al.*, 1997. Despite the advances in molecular cloning, the biosynthetic pathway from glucose to sugars found in anthracyclines is still mainly hypothetical.

Regarding the genes for deoxyhexose pathway, Madduri *et al.* (1998) have reported that a gene derived from avermectin biosynthesis cluster caused the production of hybrid anthracyclines altering the sugar moiety when transferred into an *S. peucetius* mutant. The product obtained was epirubicin, a commercially important anthracycline. In this case a hydroxy group in the daunosamine moiety was in the opposite stereochemistry due to the action of an avermectin biosynthesis gene. *S. galilaeus* has been used as the host to prepare hybrid anthracyclines using the genes derived from rhodomycin pathway from *S. purpurascens* (Niemi *et al.*, 1994), and from nogalamycin biosynthesis cluster from *S. nogalater* (Ylihonko *et al.*, 1996a). The genes for nogalamycin pathway were used to generate the hybrid anthracycline production in *S. steffisburgensis* producing

typically steffimycin (Kunnari *et al.*, 1997). Previously, biosynthesis genes for actinorhodin have been expressed in *S. galilaeus* resulting in the formation of aloesaponarin (Strohl *et al.*, 1991). These hybrid compounds were modified in the aglycone moiety.

Summary of the invention

The present invention concerns a gene cluster of *Streptomyces nogalater*, most of the genes of the cluster being derived from the deoxyhexose pathway for nogalamine and nogalose. Expressing a DNA fragment of the said region in *S. galilaeus*, which produces aclacinomycins, hybrid anthracyclines are obtained, wherein the aglycone moiety is derived from *S. galilaeus*, whereas the sugar moiety is characteristic neither to *S. nogalater* nor to *S. galilaeus*. Furthermore, when inserting the gene included in said cluster, encoding a cyclase for nogalamycin, into a suitable plasmid construction, nogalamycinone is obtained, which is the aglycone of nogalamycin. Since the stereochemistry of nogalamycin differs from most other anthracyclines, using this gene enables the preparation of C–9 stereoisomers of the anthracycline molecules.

Detailed description of the invention

The experimental procedures of the present invention are methods conventional in the art. The techniques not described in detail here are given in the manuals by Hopwood *et al.* "Genetic manipulation of Streptomyces: a laboratory manual" The John Innes Foundation, Norwich (1985) and by Sambrook *et al.* (1989) "Molecular cloning: a laboratory manual". The publications, patents and patent applications cited herein are given in the reference list in their entirety.

The present invention concerns particularly the gene cluster for nogalamycin biosynthesis (*Sno*5–cluster) causing the production of hybrid antibiotics with modifications in the sugar moiety. The invention concerns in specific the use of the genes for nogalamine/nogalose biosynthesis to generate hybrid antibiotics modified in sugar moieties. The invention also concerns the use of a specific cyclase gene included in the

gene cluster of the invention, to generate the C–9 stereoisomers of typical anthra-cyclines.

The gene cluster according to the present invention is linked to the earlier reported clusters for nogalamycin biosynthesis. The starting point of the present invention was the gene cluster for nogalamycin chromophor (International Patent Application WO 96/10581). Subsequently, we have found some genes for the deoxyhexose pathway of nogalamycin biosynthesis (Torkkell *et al.*, 1997), and a part of the fragment comprising said genes was used to clone the genes for this invention.

The biosynthesis genes for nogalamycin can be isolated from *Streptomyces* sp., particularly from *S. nogalater*, which produces nogalamycin. Species which produce nogalamycin–like anthracyclines can also be used, e.g. *S. violaceochromogenes* producing arugomycin (Kawai *et al.*, 1987), or *S. avidinii* producing avidinorubicin (Aoki *et al.*, 1991).

Genomic DNA of a *Streptomyces* strain carrying the genes for nogalamycin biosynthesis is used in preparing a genomic library. Suitable gene fragments for cloning may be obtained by any frequently digesting restriction enzyme. Typically *Sau*3AI is used. The isolated fragments could be inserted by ligation in any *Escherichia coli* vector such as a plasmid, a phagemid, a phage, or a cosmid. A cosmid vector is preferred since it enables the cloning of large DNA fragments. A cosmid vector such as pFD666 (ATCC No. 77286) is suitable for this purpose, as it enables cloning of the fragments of about 40 kb. The *Bam*HI site of pFD666, giving sticky ends to the *Sau*3AI fragments may be used for cloning. Commercially available kits may be used to pack the DNA in phage particles. Various *E. coli* strains can be used for the infection by the DNA packed. An appropriate *E. coli* strain is, e.g. XL1Blue MRF', which is deficient in several restriction systems.

Using *E. coli* as a host strain for the genomic library, hybridization is an advantageous screening strategy. The probe for hybridization may be any known fragment derived from the nogalamycin gene cluster, but a short fragment of about 1 kb derived from one

end of the biosynthetic region previously cloned is preferred. Colonies for the genomic library are transferred for filter hybridization to membranes, preferably to nylon membranes. Since the average size for a genomic DNA fragment is 40 kb, 2300 colonies gave 99.99% probability to find the expanded region for nogalamycin biosynthesis. Any method for hybridization may be used but, in particular, the DIG System (Boehringer Mannheim, GmbH, Germany) is useful. Since the probe is homologous to the hybridized DNA, it is preferable to carry out the stringent washes of hybridization at 70°C in a low salt concentration according to Boehringer Mannheim's manual "DIG System User's Guide for Filter Hybridization". At least 80% homology is suggested to be needed for a DNA fragment to bind a probe in the conditions used for washes.

Using this protocol, seven clones out of about 5000 gave positive signals, and were picked up for DNA isolation. Restriction mapping is an appropriate technique for characterizing the clones. The positive clones may be digested with convenient restriction enzymes to demonstrate the physical linkage map of the DNA fragments. The cosmid used for cloning was a shuttle cosmid replicating in both *E. coli* and *Streptomyces* sp. However, the transfer of the recombinant cosmids in *S. lividans* TK24, which is a typically used laboratory strain in cloning *Streptomyces*, resulted in deletions, and was omitted. Instead, we rather used in the expression studies the plasmid pIJ486, a high copy number *Streptomyces* plasmid. However, any plasmid being able to stably replicate in *Streptomyces* may be used for this purpose.

Two *Bgl*II fragments of one of the clones were separately inserted into pIJ486 vectors, and the two plasmids obtained were transferred into a primary host, *S. lividans* TK24. The recombinant plasmids obtained (pSY42 and pSY43), containing a 10 kb and a 7kb fragment from *S. nogalater* genomic DNA, respectively, were isolated from the primary host and further introduced into other *Streptomyces* strains by protoplast transformation. The recombinant plasmid containing the 10 kb fragment caused the production of hybrid anthracyclines in the *S. galilaeus* mutant strain H039, which endogenously produces aklavinone–rhodinose–rhodinose–rhodinose. A few other *S. galilaeus* strains (H075, H026, H063) mutated in deoxyhexose pathway for sugars in aclacinomycin were used in

transformation, and new hybrid compounds were obtained. Since the structure of nogalamycin is almost unique among anthracyclines, the plasmids could be transferred to other anthracycline-producing strains, such as *S. peucetius*, which produces daunomycin, and *S. purpurascens*, which produces rhodomycins, to modify the structures of the characteristic antibiotics.

As the cloned cluster was linked to nogalamycin biosynthesis region already known, its ability to generate the modification in sugar moiety suggested the presence of the genes for deoxyhexose pathway. However, sequencing is necessary to deduce the function of the genes in the cluster cloned. The DNA fragments of 10 kb and 7 kb were further inserted into the plasmid pSL1190 for subcloning. Sequencing strategies such as a deletion set of the DNA fragments, shotgun cloning or primer walking could be used, but we prefer to use restriction fragments for subcloning. Using ABI PRISM system (Perkin-Elmer) for sequencing it is possible to get 500 to 700 bases per one reaction, which means that about 1 kb fragments sharing overlapping bases are needed for sequencing. For this purpose, 27 subclones were constructed.

Sequencing of the flanked *Bgl*II fragments consisting of about 16000 bp revealed 15 complete ORFs. The sequence analysis can be made by any computer based program, such as GCG (Madison, Wisconsin, USA) package. According to the present invention the putative gene functions as deduced from the sequence homology of those available in the libraries are

aminotransferase (*sno*gI), not completed

1. dTDP-glucose synthase (*sno*gJ)
2. aminomethyl transferase (*sno*gA)
3. polyketide cyclase, (*sno*aM)
4. a gene of deoxyhexose pathway, unknown (*sno*gN)
5. hydroxylase, (*sno*aG)
6. dTDP-4-dehydrorhamnose reductase (*sno*gC)
7. dTDP-glucose 4,6-dehydratase (*sno*gK)
8. NAME cyclase (*sno*aL)
9. unknown (*sno*K)

10. glycosyl transferase, GTF (*sno*gD)
11. unknown (*sno*W)
12. glycosyl transferase, GTF (*sno*gE)
13. unknown (*sno*L)
14. unknown (*sno*O)
15. C–7 ketoreductase (*sno*aF)

unknown (*sno*N), not completed

Gene designations: **g** means that the gene involved in biosynthesis of the glycosidic proportion including glycosyl transferases, whereas **a** points out that the gene is needed for the formation of the aglycone moiety.

Considering the proposed biosynthetic pathway for nogalamycin shown in Fig 3. we are able to cause several changes for the structures of antibiotics by the genes identified, including *sno*aL, responsible for the cyclization of the fourth ring of the aglycone moiety while determining the stereochemistry of the anthracyclinone, and the genes affecting the formation of nogalamine and nogalose (*sno*gJ, *sno*gK, *sno*gN, *sno*gC, *sno*gA), and, in addition, the genes responsible for joining the sugar residues to the aglycone moiety (*sno*gD and *sno*gE).

These genes could be separately inserted in a vector using suitable restriction sites, or by amplifying the genes by PCR. The fragments may contain an intrinsic promoter, or a promoter may be separately cloned. It is advantageous to use a vector carrying a promoter to allow expression of the genes in a *Streptomyces* strain. The plasmid pIJE486 contains a promoter *erm*E for erythromycin resistance gene, allowing constitutive expression of the genes inserted in a correct orientation. Special attention is drawn to the gene encoding a cyclase for the aliphatic ring, but any gene of said cluster may be expressed in *Streptomyces* hosts. The said cyclase converts the stereochemistry at C9 of auramycinone in TK24, if inserted into the plasmid possessing the other genes for auramycinone biosynthesis, except the cyclase responsible for the typical stereochemistry of anthracyclines.

Streptomyces strains, in particular *S. galilaeus*, carrying the recombinant plasmids are cultivated in media wherein antibiotics are produced. The hybrid compounds are extracted with organic solvents from the culture broth, and the compounds are separated and purified using chromatographic techniques.

According to this invention *S. galilaeus* H039 carrying the plasmid pSY42 and designated as H039/pSY42 produces aklavinone–4'–epi–2–deoxyfucose in E1 medium supplemented with thiostrepton to give selection pressure for the plasmid containing strains.

S. lividans TK24 carrying the plasmid pSY15c containing the genes for the nogalamycin chromophor and the genes for a cyclase (*sno*aL) and a ketoreductase (*sno*aF), was cultivated in E1 medium supplemented with thiostrepton. The compound 9–epi–auramycinone was produced, and this structure is now called nogalamycinone. Any DNA fragment of the invention subcloned from a 17 kb nogalamycin biosynthesis region can be inserted in a vector replicating in *Streptomyces*, and the products may be produced by fermentation of the plasmid containing strains.

Brief description of the drawings

Fig. 1 shows the structures of nogalamycin, daunomycin and aclacinomycin.

Fig. 2 is a diagram of the gene cluster (*Sno5*) of the invention for nogalamycin biosynthesis.

Fig. 3 describes the proposed biosynthesis pathway for nogalamycin.

Fig. 4 shows a diagram of the plasmid pSY15c. The genes *sno*aL (**aL**) and *sno*aF (**aF**) shown black are inserted in the plasmid pSY15 to give pSY15c. **aL** represents a cyclase *sno*aL and **aF** is for C–7 ketoreductase *sno*aF. pSY15 (WO 96/10581) generates the production of a tricyclic intermediate for nogalamycin biosynthesis in *S. lividans*. The abbreviations **a1**, **a2** and **a3** refer to the

genes *sno*a1, *sno*a2 and *sno*a3, respectively, for minimal PKS. **rA** is the *snor*A gene for an activator, **aB** is the *sno*aB gene for oxygenase, **aC** is the *sno*aC gene for methylase, **aD** is the *sno*aD gene for polyketide ketoreductase and **aE** is the *sno*aE for aromatase. **gF** (the *sno*gF gene) and **gG** (the *sno*gG gene) involved in the deoxyhexose pathway are not functional in the construct. **aph** is an aminoglycoside phosphotransferase gene, and **tsr** is a thiostreptone resistance gene.

Examples to further illustrate the invention are given hereafter.

EXPERIMENTAL

Materials used

Restriction enzymes used were purchased from Promega (Madison, Wisconsin, USA) or Boehringer Mannheim (Germany), and alkaline phosphatase from Boehringer Mann-heim, and used according to the manufacturers' instructions. Proteinase K was purchased from Promega and lysozyme from Sigma (St. Louis, USA). Hybond™-N nylon membranes used in hybridization were purchased from Amersham (Buckinghamshire, England), DIG DNA Labelling Kit and DIG Luminescent Detection Kit from Boehringer Mannheim. Qiaquick Gel Extraction Kit from Qiagen (Hilden, Germany) was used for isolating DNA from agarose.

Bacterial strains and their use

- *Escherichia coli* XL1 Blue MRF' (Stratagene, La Jolla, CA) was used for cloning.
- *Streptomyces nogalater* ATCC 27451; the gene cluster of nogalamycin biosynthesis was cloned from this strain.

The host strains to express the genes cloned were:

- *Streptomyces lividans* TK24, also used as a primary host to clone DNA propagated in *E. coli*. The strain was provided by prof. Sir David Hopwood, John Innes Centre, UK.
- *Streptomyces galilaeus* H039, produces aklavinone-rhodinose-rhodinose-rhodinose
- *Streptomyces galilaeus* H026, produces aclacinomycin N, ACMN, (aklavinone-rhodosamine-2-deoxyfucose-rhodinose)

– *Streptomyces galilaeus* H063, produces aklavinone

– *Streptomyces galilaeus* H075, produces aklavinone–rhodosamine–2–deoxyfucose–2–deoxyfucose

The detailed description of the mutants H039 and H026 is given in Ylihonko *et al.* (1994) and of H075 in the FI patent application No. 981062 (Ylihonko *et al.*, 1998). H063 has not been described in the literature but it was obtained by NTG mutagenesis of *S. galilaeus*, and selected to be used as the host strain in the hybrid compound production, as it accumulates aklavinone without any sugar residues.

Plasmids

E. coli – Streptomyces shuttle cosmid pFD666 (ATCC 77286) was used for cloning the chromosomal DNA. *E. coli* cloning vectors pSL1190 (Pharmacia) and pUC19 were used for preparing the subclones.

pIJ486 is a high copy plasmid vector provided by prof. Sir David Hopwood, John Innes Centre, UK (Ward *et al.*, 1986)

pIJE486 is a vector containing *erm*E gene in the polylinker of pIJ486 (Bibb *et al.*, 1985).

pSY15 is a pIJ486 based plasmid construct, wherein the genes of polyketide pathway for nogalamycin biosynthesis were cloned (Ylihonko *et al.*, 1996a).

Nutrient media and solutions

For cultivation of *S. nogalater* for total DNA isolation TSB medium was used. Lysozyme solution (0.3 M sucrose, 25 mM Tris, pH 8 and 25 mM EDTA pH 8) was used in isolation of total DNA. TE buffer (10 mM Tris, pH 8.0 and 1mM EDTA) was used to dissolve the DNA.

TRYPTONE–SOYA BROTH (TSB)

Per litre: Oxoid Tryptone Soya Broth powder 30 g.

ISP4

Bacto ISP–medium 4, Difco; 37 g/l.

E1 Per litre in tap water:

glucose	20	g
soluble starch	20	g
Farmamedia	5	g
yeast extract	2.5	g
$K_2HPO_4 \bullet 3H_2O$	1.3	g
$MgSO_4 \bullet 7H_20$	1	g
NaCl	3	g
$CaCO_3$	3	g

pH adjusted to 7.4 before autoclaving

General methods

NMR data was collected with a JEOL JNM–GX 400 spectrometer at the ambient temperature. ^1H and ^{13}C NMR samples were internally referenced to TMS.

The anthracycline metabolites were detected by HPLC (LaChrom, Merck Hitachi, pump L–7100, detector L–7400 and integrator D–7500) using a LiChroCART RP–18 column (4.6x250mm). Acetonitrile:potassium hydrogen phosphate buffer (60 mM, pH 3.0 adjusted with citric acid) was used as the mobile phase. Gradient system starting from 65% to 30% of potassium dihydrogen phosphate buffer was used to separate the compounds. The flow rate was 1 ml/min and the detection was effected at 430 nm.

ISP4 plates supplemented with thiostrepton (50 μg/ml) were used to maintain the plasmid carrying cultures.

Example 1. Cloning the gene cluster for nogalamycin biosynthesis

1.1 Cosmid library

For the isolation of total DNA, *Streptomyces nogalater* (ATCC 27451) was grown for three days in 50 ml of TSB medium supplemented with 0.5% of glycine. The cells were harvested by centrifuging for 15 min at 3900 x g in 12 ml Falcon tubes, and the

cells were stored at –20°C. Cells from a 12 ml sample of the culture were used to isolate the DNA. 5 ml of lysozyme solution containing 5 mg of lysozyme/ml was added onto the cells, incubated for 20 min at 37°C. 500 μl of 10% SDS containing 0.7 mg of proteinase K was added onto the cells and incubated for 80 min at 62°C, another 500 μl of 10% SDS containing 0.7 mg of proteinase K was added, and incubation was continued for 60 min. The sample was chilled on ice and 600 μl of 3M NaAc, pH 5.8 were added, and the mixture was extracted with equilibrated phenol (Sigma). The phases were separated by centrifuging at 1400 x g for 10 min. The DNA was precipitated from the water phase with equal volume of isopropanol to spool with a glass rod, and washed by dipping to 70% ethanol, air dried and dissolved in 500 μl of TE–buffer.

The chromosomal DNA was partially digested with *Sau*3AI. The DNA fragments were separated by agarose gel electrophoresis, and the fragments of 30 to 50 kb were cut from the 0.3% low gelling temperature SeaPlaque® agarose. The DNA bands were isolated from the gel by heating to 65°C, extracting with equal volume of equilibrated phenol, and the phases were separated by centrifuging for 15 min at 2500 x g. The phenol phase was extracted with TE buffer, centrifuged and the water phases were pooled. The DNA was precipitated by adding 0.1 volumes of NaAc, pH 5.8 and 2 volumes of ethanol at –20°C for 30 min, centrifuged for 30 min at 15 000 rpm in Sorvall RC5C centrifuge using SS–34 rotor with adapters for 10 ml tubes. The pellet was air dried and dissolved in 20 μl of TE buffer. The isolated fragments were ligated to pFD666 cosmid vector digested with *Bam*HI and dephosphorylated. The DNA was packed into phage particles, and infected to *E. coli* using Gigapack® III XL Packing Extract Kit according to the manufacturer's instructions.

1.2 Identification of the clones by hybridization

The infected cells were grown on LB plates containing 50 μg/ml kanamycin and transferred to Hybond™–N nylon membranes (Amersham). The membranes were handled according to the protocol described in Boehringer Mannheim's manual "The DIG System User's Guide for Filter Hybridization". The probe used to screen the colonies for an expanded nogalamycin gene cluster was a 1.07 kb *Sac*I fragment from the cluster described earlier (Torkkell *et al.*, 1997). The plasmid carrying the probe was

digested with *Sac*I, and the fragment was separated from the vector by agarose gel electrophoresis and isolated from the gel using Qiaquick Gel Extraction Kit (Qiagen). The probe was labelled by digoxygenin using random prime labelling system according to Boehringer Mannheim's manual "The DIG System User's Guide for Filter Hybridization". 5000 colonies were screened by hybridization at 70°C using the probe described. Positive colonies were detected using DIG Luminescent Detection Kit (Boehringer Mannheim). Seven colonies gave a positive signal. Cosmids from the positive clones were isolated from a 5ml culture by alkaline lysis method. Restriction analysis showed that the cloned fragments overlapped each other representing at least 60 kb of the continuous DNA. The positive clones obtained were designated as pFD*Sno*1 to pFD*Sno*7.

1.3. Subcloning the fragments for sequencing

Clone No. 5, designated as pFD*Sno*5, was digested with *Bgl*II, and for subcloning two fragments of about 10 kb and 7 kb were isolated and ligated to pSL1190 digested with *Bgl*II and dephosphorylated. The plasmids obtained were named as pSn42 and pSn43, respectively. These two fragments cover the DNA region flanked to the previously characterized area of nogalamycin biosynthesis cluster. To determine the nucleotide sequence of the whole 17 kb region cloned in pSn42 and pSn43 the convenient restriction sites were used to subclone the fragments to the vector pUC19 or pSL1190 giving 16 subclones from the insert of pSn42 and 11 subclones of pSn43.

E. coli XL1 Blue MRF' cells were cultivated overnight at 37 °C in 5 ml of LB-medium supplemented with 50 μg/ml of ampicillin. To isolate plasmids for sequencing reactions Wizard Plus Minipreps DNA Purification System kit of Promega, or Biometra silica spin plasmid miniprep kit of Biomedizinische Analytik Gmbh were used according to the manufacturers' instructions.

DNA sequencing was performed using the automatic ABI DNA sequenator (Perkin-Elmer) according to the manufacturer's instructions.

1.4 Sequence analysis and the deduced functions of the genes

Sequence analyses were effected using the GCG sequence analysis software package (Version 8; Genetics Computer Group, Madison, Wisconsin, USA). The translation table was modified to accept also GTG as a start codon. Codon usage was analysed using published data (Wright and Bibb 1992).

According to the CODONPREFERENCE program the sequenced DNA fragment (SEQ ID NO:1) contained 15 complete open reading frames (ORFs), and the 5' end of other two ORFs in the both ends of the fragment according to the invention. The functions of the genes were concluded by comparing the amino acid sequences translated from their base sequences to the known protein sequences in the data banks. The results are shown in Table 1. The positions given refer to the appended sequence listing. The amino acid sequences of the peptides are given in SEQ ID NO:2 to SEQ ID NO:18.

Table 1

Gene	Position	Amino acids (SEQ ID NO)	Deduced function	Remarks
*snog*I	-1027 compl	>342 (2)	aminotransferase	5' end
*snog*J	1192-2073	293 (3)	dTDP-glucose synthase	
*snog*A	2106-2822 compl	238 (4)	aminomethyl transferase	
*snoa*M	2826-3800 compl	324 (5)	a polyketide cyclase	
*snog*N	3799-5025	408 (6)	*dnr*Q homology (Otten *et al.*, 1995), unknown	
*snoa*G	5088-6356	422 (7)	hydroxylase	
*snog*C	6334-7209 compl	291 (8)	dTDP-4-dehydrorhamnose reductase	
*snog*K	7245-8297 compl	350 (9)	dTDP-glucose-4,6-de-hydratase	
*snoa*L	8537-8941	134 (10)	NAME cyclase (nogalonic acid methyl ester)	
*sno*K	8992-9699	235 (11)	unknown	
*snog*D	9745-10917 compl	390 (12)	glycosyl transferase	
*sno*W	11057-11884	275 (13)	unknown	
*snog*E	11928-*	>424 (14)	glycosyl transferase	
*sno*L	13335-13754 compl	139 (15)	unknown	
*sno*O	13974-14441	155 (16)	homologous to *mtm*X of mithramycin cluster	
*snoa*F	14532-15377	281 (17)	C-7 ketoreductase analog-ous to aklaviketone keto-reductase	
*sno*N	15450-	>190 (18)	unknown	5' end

*: nucleotide sequence of about 100 bp, not known

1.5 Expression cloning

The 10 kb *Bgl*II fragment from pFD*Sno*5 was cloned into the plasmid pIJ486 and the plasmid obtained was named as pSY42. Correspondingly, the 7 kb *Bgl*II fragment from pFD*Sno*5 was cloned into the plasmid pIJE486, and the plasmid pSY43 was obtained. Plasmid pSY42 was introduced into *S. lividans* strain TK24 by protoplast transformation, isolated from it and further introduced into *S. galilaeus* mutant H039, and after propagation in H039, transferred to other *S. galilaeus* mutants blocked in the deoxyhexose pathway for characteristic sugars of aclacinomycins (H075, H026, and H063). E1 medium was used for anthracycline production, and the products were extracted from the culture with toluene:methanol (1:1) at pH 7. Anthracycline metabolites were analyzed by HPLC. The products of the mutants H039, H026, H063 and H075 carrying pSY42 differed from those obtained by the mutants without the plasmid.

According to the sequence analysis pSY42 contained a cyclase designated as NAMEC (nogalonic acid methyl ester cyclase), and in pSY43 a ketoreductase gene was identified. Expression constructions were prepared which contained all the genes needed for the formation of nogalamycin aglycone. A 1.4 kb *Bam*HI–*Sac*I fragment from pSY42 (containing NAMEC) and a 1.1 kb *Mlu*I–*Kpn*I fragment from pSY43 carrying the gene for a ketoreductase of C–7 keto group were ligated to pSY15 linearized by *Sac*I, to form the plasmid pSY15c (Fig. 4). Plasmid pSY15c was introduced into *S. lividans* TK24, and the strain TK24/pSY15c was cultivated in E1 medium supplemented with thiostrepton. An aglycone compound was produced, and this structure is now called nogalamycinone.

Example 2. Compounds generated by the *sno*5–cluster

2.1 Production and purification of the products derived from H039/pSY42 and TK24/pSY15c

The seed culture, 180 ml of E1 culture of the plasmid containing strain, H039/pSY42 or TK24/pSY15c, was obtained by cultivating the strain in three 250 ml Erlenmeyer flasks containing 60 ml of E1 medium supplemented with thiostrepton (5 μg/ml) for four days at 30°C, 330 rpm. The combined culture broths (180 ml) were used to inoculate 13 l of

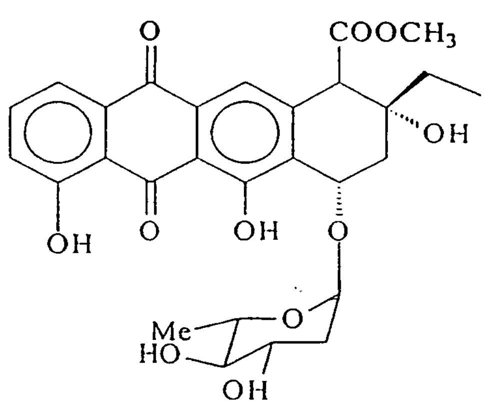
E1 medium in a fermentor (Biostat E). Fermentation was carried out for seven days at 28°C (330 rpm, aeration: 450 l/min).

The cells were harvested by centrifuging. 2.6 l of methanol was used to break the bacterial cells and to extract anthracycline metabolites accumulated. The anthracycline metabolites were extracted using 2 l of dichloromethane at pH 6. The organic layer was evaporated to dryness. The viscous residue was flashed through a polyamide (11) column using water:methanol from 1:9 to 0:10 as the eluent. Pooled fractions containing the compounds were further purified on a Merck–Hitachi HPLC using preparative reversed phase column (LichroCART RP–18, 5 μm) with mobile phase acetonitrile:1 % AcOH in water (1:1). Evaporation of acetonitrile gave pure products as yellow powders dried under vacuum.

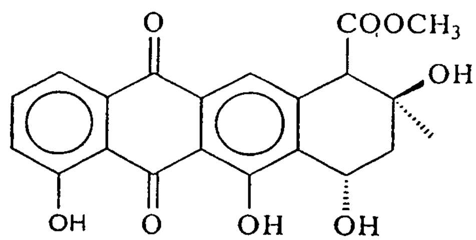
2.2 Structural elucidation of the compounds derived from H039/pSY42 and from TK24/pSY15c

NMR analysis included NON, BMC, NOE, DEPT and HMBC techniques. Protons were assignated using NOESY and 2D pTOCSY techniques and carbons using DEPT and HMBC techniques.

As deduced from the data given in Tables 2 and 3, the structures revealed were aklavinone–4'–epi–2–deoxyfucose from the culture of H039/pSY42, and 9–epi–auramycinone (=nogalamycinone) from the culture of TK24/pSY15c. The chemical structures of the compounds are shown below in Formula I and Formula II, respectively.



(I)



(II)

Deposited microorganisms

The following microorganisms were deposited according to the Budapest Treaty at Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Mascheroder Weg 1b, D–38124 Braunschweig, Germany.

Microorganism	**Accession number**	**Date of deposit**
S. lividans TK24/pSY42 carrying the plasmid pSY42	DSM 12451	14 October 1998
S. lividans TK24/pSY43 carrying the plasmid pSY43	DSM 12452	14 October 1998

Table 2. ^{1}H and ^{13}C assignations of the compound aklavinone–4'–epi–2–deoxyfucose (Formula I).

Site	^{1}H	^{13}C
1	7.74, 1H, dd, 7.5, 1.3	120.1
2	7.68, 1H, dd,. 8.4, 7.5	137.3
3	7.27, 1H, dd, 8.3, 1.3	124.6
4	–	161.9
4–OH	11.70, 1H, s	–
4a	–	115.4
5	–	192.3
5a	–	114.4
6	–	162.4
6–OH	12.46, 1H, s	–
6a	–	130.9
7	5.18, 1H, dd, 4.3, 3.1	71.3
8A	2.51, 1H, dd, 15.0, 4.3	33.9
8B	2.32, 1H, dd, 15.0, 3.1	–
9	–	72.1
9–OH	4.58, 1H, s	–
10	4.02, 1H, s	56.9
10a	–	142.4
11	7.40, 1H, s	120.8
11a	–	133.1
12	–	180.7
12a	–	132.6
13A	1.73, 1H, dq, 14.2, 7.4	32.0
13B	1.51, 1H, dq, 14.2, 7.4	–
14	1.10, 3H, t, 7.4	6.7
15	–	171.1
16	3.69, 3H, s	52.5
1´	5.41, 1H, d, 3.5	101.7
2´a	1.75, 1H, ddd, 12.8, 11.2, 3.4	37.7
2´e	2.19, 1H, dd, 12.8, 5.3	–
3´	3.71, 1H, ddd, 12.0, 9.0, 5.3	69.0
4´	3.14, 1H, dd, 9.1, 9.0	78.1
5´	3.88, 1H, dq, 9.1, 6.2	68.8
6´	1.36, 3H, d, 6.2	17.6

Table 3. ^{1}H and ^{13}C assignations of 9-epi-auramycinone (Formula II).

Site	^{1}H	^{13}C
1	7.76, 1H, dd, 7.5, 1.2	119.8
2	7.67, 1H, dd, 8.3, 7, 5	137.4
3	7.28, 1H, dd, 8.3, 1.2	124.8
4	–	162.5
4–OH	11.86, 1H, s	–
4a	–	115.6
5	–	192.7
5a	–	114.6
6	–	160.9
6–OH	12.76, 1H, s	–
6a	–	134.1
7	5.40, 1H, t, 7.0	64.0
8A	2.66, 1H, dd, 13.9, 7.0	40.9
8B	1.89, 1H, dd, 13.9, 7.1	–
9	–	70.5
9–OH	3.49, 1H, brs	–
10	3.93, 1H, d, 0.8	56.0
10a	–	142.1
11	7.51, 1H, d, 0.8	120.1
11a	–	133.3
12	–	180.9
12a	–	132.1
13	1.44, 3H, s	28.7
14	–	173.0
15	3.90, 3H, s	52.6

References

Aoki, M., Shirai, H., Nakayama, N., Itezono, Y., Mori, Mitsuko, Satoh, T., Ohishima, S., Watanabe, J., and Yokose, K. 1991. Structural studies on avidinorubicin, a novel anthracycline with platelet aggregation inhibitory activity. J. Antibiot. **44:** 635–645.

Bibb, M.J., Janssen, G.R., and Ward, J.M. 1985. Cloning and analysis of the promoter region of the erythromycin resistance gene (*ErmE*) of *Streptomyces erythraeus*. Gene **38**: 215–226.

Bhuyan, B.K., and Dietz, A. 1965. Fermentation, taxonomic and biological studies of nogalamycin. Antimicrob. Ag. Chemother. 1965: 836–844.

Bhuyan, B.K., and Smith C.G. 1975. In: Antineoplastic and immunosuppressive agents, part II, pp. 623–632. Sartorell, A.C., and Johns, D.G. (eds), Springer–Verlag, Berlin.

Fujii, I., and Ebizuka, Y. 1997. Anthracycline biosynthesis in *Streptomyces galilaeus*. Chem. Rev. **97**: 7: 2511–2523.

Kawai, H., Hayakawa, Y., Nakagawa, M., Furihata, K., Shimazu, A., Seto, H. and Otake, N. 1987. Arugomycin, a new anthracycline antibiotic I. Taxonomy, fermentation, isolation and physico–chemical properties. J. Antibiot. **19:** 1266–1271.

Kunnari, T., Tuikkanen, J., Hautala, A., Hakala, J., Ylihonko, K., and Mäntsälä, P. 1997. Isolation and characterization of 8–demethoxy steffimycins and generation of 2,8–demethoxy steffimycins in *Streptomyces steffisburgensis* by the nogalamycin biosynthesis genes. J. Antibiot. **50:** 496–501.

Lomovskaya, N., Doi–Katayama, Y., Filippini, S., Nastro, C., Fonstein, L:, Gallo, M., Colombo, A.L. and Hutchinson, C.R. 1998. The *Streptomyces peucetius dpsY* and *dnrX* genes govern early and late steps of daunorubicin and doxorubicin biosynthesis. J. Bacteriol. **180**: 9: 2379–2386.

Madduri, K., Kennedy, J., Rivola,G., Inventi–Solari, A., Filippini, S., Zanuso, G., Colombo, A.L., Gewain, K.M., Occi, J.L., MacNeil, D.J., and Hutchinson, C.R. 1998. Production of antitumor drug epirubicin (4'–epidoxorubicin) and its precursor by a genetically engineered strain of *S. peucetius*. Nature Biotech. **16**: 69–74.

Niemi, J., Ylihonko, K., Hakala, J., Kopio, A., Pärssinen, R., and Mäntsälä, P. 1994. Hybrid anthracycline antibiotics: production of new anthracyclines by cloned genes from *Streptomyces purpurascens* in *Streptomyces galilaeus*. Microbiol. **140:** 1351–1358.

Otten, S.L., Liu, X., Ferguson, J., and Hutchinson C.R. 1995. Cloning and characterization of the *Streptomyces peucetius dnrQS* genes encoding a daunosamine biosyntesis enzyme and a glycosyl transferase involved in daunorubicin biosynthesis. J. Bacteriol. **177**: 6688–6692.

Rajgarhia, V.B., and Strohl, W.R. 1997. Minimal *Streptomyces* sp. strain C5 daunorubicin polyketide biosynthesis genes required for aklanonic acid biosynthesis. J. Bacteriol. **179**: 8: 2690–2696.

Strohl, W.R., Bartel, P.L., Li, Y., Connors, N.C., and Woodman, R.H. 1991. Expression of polyketide biosynthesis and regulatory genes in heterologous streptomycetes. J. Ind. Microbiol. **7**: 3: 163–174.

Torkkell, S., Ylihonko, K., Hakala, J., Skurnik, M., and Mäntsälä, P. 1997. Characterization of *Streptomyces nogalater* genes encoding enzymes involved in glycosylation steps in nogalamycin biosynthesis. Mol. Gen. Genet. **256**: 203–209.

Ward, J.M., Janssen, G.R., Kieser, T., Bibb, M.J., Buttner, M.J., and Bibb, M.J. 1986. Construction and characterization of a series of multicopy promoter–probe plasmid vectors for Streptomyces using the aminoglycoside phosphotransferase from Tn5 as indicator. Mol. Gen. Genet. **203**: 468–478.

Wiley, P.F., Kelly, R.B., Caron, E.L., Wiley, V.H., Johnson, J.H., MacKellar, F.A. and Mizsak, S.A. 1977. Structure of nogalamycin. J. Am. Chem. Soc. **99**: 542–549.

Wiley, P.F., Elrod, D.W., and Marshall, V.P. 1978. Biosynthesis of the anthracycline antibiotics nogalamycin and steffimycin. B. J. Org. Chem. **43**: 3457–3461.

Wright, F., and Bibb, M.J. 1992. Codon usage in the G+C –rich *Streptomyces* genome. Gene **113**: 55–65.

Ylihonko, K., Hakala, J., Niemi, J., Lundell, J., and Mäntsälä, P. 1994. Isolation and characterization of aclacinomycin A–non–producing *Streptomyces galilaeus* (ATCC 31615) mutants. Microbiol. **140**: 1359–1365.

Ylihonko, K., Hakala, J., and Mäntsälä, P. 1995. Process for producing anthracyclines and intermediates thereof. WO 96/10581.

Ylihonko, K., Hakala, J., Kunnari, T., and Mäntsälä, P. 1996a. Production of hybrid anthracycline antibiotics by heterologous expression of *Streptomyces nogalater* nogalamycin biosynthesis genes. Microbiol. **142**: 1965–1972.

Ylihonko, K., Tuikkanen, J., Jussila, S., Cong, L., and Mäntsälä, P. 1996b. A gene cluster involved in nogalamycin biosynthesis from *Streptomyces nogalater*: sequence analysis and complementation of early–block mutations in the anthracycline pathway. Mol. Gen. Genet. **251**: 113–120.

Ylihonko, K., Hakala, J., and Kunnari, T. 1998. Hybrid anthracyclines from genetically engineered *Streptomyces galilaeus* strains. No. 981062.

Yoshida, M., Fujioka, A., Nakano, K., Yuasa, C., Toko, T., Takeda, S., and Unemi, N. 1996. Activity of menogaril against various malignant lymphoma cell lines and a human lymphoma xenograft in mice. Anticancer Res. **16**: 2875–2879.